# OC3140
# HW/Lab 8 Analysis of Variance

The follow table lists temperature recorded at Carmel, Monterey, Seaside, and Salinas.

| | Carmel | Monterey | Seaside | Salinas |
|---|---|---|---|---|
| 6:00 AM | 58 | 58 | 57.6 | 66.6 |
| 8:00 AM | 59 | 58.8 | 58.5 | 67 |
| 10:00 AM | 60 | 61 | 60 | 69 |
| 12:00 AM | 61 | 61 | 61 | 70 |
| 2:00 PM | 62 | 61.5 | 61.7 | 71 |
| 4:00 PM | 62.5 | 62 | 62.5 | 71.5 |
| 6:00 PM | 62 | 61.5 | 62 | 71 |
| 8:00 PM | 61 | 61 | 61.5 | 70 |

1. Determine if the mean temperatures at these four cities are the same (One-Way). A significance level of 5% is used

**Solution** see Ch.8 p2-p7

Fill the table as

| | Carmel | Monterey | Seaside | Salinas | Grand |
|---|---|---|---|---|---|
| 6:00 AM | 58 | 58 | 57.6 | 66.6 | |
| 8:00 AM | 59 | 58.8 | 58.5 | 67 | |
| 10:00 AM | 60 | 61 | 60 | 69 | |
| 12:00 AM | 61 | 61 | 61 | 70 | |
| 2:00 PM | 62 | 61.5 | 61.7 | 71 | |
| 4:00 PM | 62.5 | 62 | 62.5 | 71.5 | |
| 6:00 PM | 62 | 61.5 | 62 | 71 | |
| 8:00 PM | 61 | 61 | 61.5 | 70 | |
| sum | 485.5 | 484.8 | 484.8 | 556.1 | 2011.2 |
| mean | 60.6875 | 60.6 | 60.6 | 69.5125 | 62.85 |
| Variance | 2.4955 | 2.0086 | 3.0743 | 3.4155 | |

n=8, p=4, N=n*p=32.

### Method 1 (One-factor ANOVA):

$$s_W^2 = \frac{\sum_{i=1}^{p} s_i^2}{p} = 2.7485, \quad \overline{X} = \frac{G}{n*p} = 62.85$$

$$s_B^2 = \frac{n\sum_{i=1}^{p}\left(\overline{x}_i - \overline{X}\right)^2}{p-1} = 157.84, \quad F = \frac{s_B^2}{s_W^2} = \frac{157.84}{2.7485} = 57.43,$$

**Method 2 (Partitioning of Variability):**

$$SST = \sum_{i=1}^{p}\sum_{j=1}^{n_i} x_{i,j}^2 - \frac{G^2}{N} = 550.48, \quad SSB = \frac{1}{n}\sum_{i=1}^{p} T_i^2 - \frac{G^2}{N} = 473.5225$$

$$SSW = SST \text{ - } SSB = 76.9575, \quad MSB = \frac{SSB}{p-1} = 157.8408$$

$$MSW = \frac{SSW}{p(n-1)} = 2.7485, \quad F = \frac{MSB}{MSW} = 57.43$$

using the F table (Ch.5 p21-p22), we have

$$F(a = 0.05, d.f. = 3,28) = 2.95 .$$

Since

$$2.95 < F(= 57.43),$$

we conclude that there is at least one pair of cities has different mean temperature at 5 % level.

2. If there are different mean temperatures, which two cities have different mean temperatures?

**Solution:** This is **pair-wise** comparison, see Ch.8 p7-p10

**Method 1 (Fisher's Least Significant Difference Method):**

- From above we know: $s_w^2 = 2.7485$
- Computer $LSD$ (n=8, p=4, N=32):

$$LSD = t_{a/2}\sqrt{\frac{2s_w^2}{n}} = 2.048\sqrt{\frac{2*2.7485}{8}} = 1.6976$$

- Pair-wise Comparison:

$$|\bar{x}_{ca} - \bar{x}_{mo}| = 0.0875 < 1.6976 = LSD$$

$$|\bar{x}_{ca} - \bar{x}_{se}| = 0.0875 < 1.6976 = LSD$$

$$|\bar{x}_{ca} - \bar{x}_{sa}| = 8.825 > 1.6976 = LSD$$

$$|\bar{x}_{mo} - \bar{x}_{se}| = 0 < 1.6976 = LSD$$

$$|\bar{x}_{mo} - \bar{x}_{sa}| = 8.9125 > 1.6976 = LSD$$

$$|\bar{x}_{se} - \bar{x}_{sa}| = 8.9125 > 1.6976 = LSD$$

The mean temperature at Salinas city has significance difference with other cities.

**Method 2 (Tukey's Method):**

- $a = 0.05$, and $p = 4$ and $N - p = 28$, the $q$ value from the table (Ch8 Appendix 6!

- $Trange = \frac{q}{\sqrt{n}}\sqrt{s_w^2} = 2.2273$

-

$$|\bar{x}_{ca} - \bar{x}_{mo}| = 0.0875 < 2.2273 = Trange$$
$$|\bar{x}_{ca} - \bar{x}_{se}| = 0.0875 < 2.2273 = Trange$$
$$|\bar{x}_{ca} - \bar{x}_{sa}| = 8.825 > 2.2273 = Trange$$
$$|\bar{x}_{mo} - \bar{x}_{se}| = 0 < 2.2273 = Trange$$
$$|\bar{x}_{mo} - \bar{x}_{sa}| = 8.9125 > 2.2273 = Trange$$
$$|\bar{x}_{se} - \bar{x}_{sa}| = 8.9125 > 2.2273 = Trange$$

The mean temperature at Salinas city has significance difference with other cities.

**Method 3 (Scheffs's Method):**

- $Sreage = \sqrt{(p-1)F_{a,p-1,N-p}} * \sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right) * s_w^2} = \sqrt{3*2.95} * \sqrt{\frac{2}{8}*2.7485} = 2.46$

-

$$|\bar{x}_{ca} - \bar{x}_{mo}| = 0.0875 < 2.466 = Srange$$
$$|\bar{x}_{ca} - \bar{x}_{se}| = 0.0875 < 2.466 = Srange$$
$$|\bar{x}_{ca} - \bar{x}_{sa}| = 8.825 > 2.466 = Srange$$
$$|\bar{x}_{mo} - \bar{x}_{se}| = 0 < 2.466 = Srange$$
$$|\bar{x}_{mo} - \bar{x}_{sa}| = 8.9125 > 2.466 = Srange$$
$$|\bar{x}_{se} - \bar{x}_{sa}| = 8.9125 > 2.466 = Srange$$

The mean temperature at Salinas city has significance difference with other cities.

3. Do Carmel, Monterey and Seaside cities have the same average temperature both in location and time (Two-factor)? A significance level of 5% is used

**Solution:** This is Two-factor (or Two-way) ANOVA (see Ch8, p10-p14)

| | Carmel | Monterey | Seaside | Total | Mean |
|---|---|---|---|---|---|
| 6:00 AM | 58 | 58 | 57.6 | 173.6 | 57.8667 |
| 8:00 AM | 59 | 58.8 | 58.5 | 176.3 | 58.7667 |
| 10:00 AM | 60 | 61 | 60 | 181 | 60.3333 |
| 12:00 AM | 61 | 61 | 61 | 183 | 61 |
| 2:00 PM | 62 | 61.5 | 61.7 | 185.2 | 61.7333 |
| 4:00 PM | 62.5 | 62 | 62.5 | 187 | 62.3333 |
| 6:00 PM | 62 | 61.5 | 62 | 185.5 | 61.8333 |
| 8:00 PM | 61 | 61 | 61.5 | 183.5 | 61.1667 |
| Total | 485.5 | 484.8 | 484.8 | 1455.1 | |
| mean | 60.6875 | 60.6 | 60.6 | | 60.6292 |

The terms are computed in the following.

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}^2 - \frac{G^2}{I * J} = 88274.59 - 888221.5 = 53.09$$

$$SSA = \sum_{i=1}^{I} \frac{T_{i.}^2}{J} - \frac{G^2}{I * J} = 88273.06 - 88221.5 = 51.56$$

$$SSB = \sum_{j=1}^{J} \frac{T_{.j}^2}{I} - \frac{G^2}{I * J} = 88221.54 - 88221.5 = 0.04$$

$$SSE = SST - SSA - SSB = 53.09 - 51.56 - 0.04 = 1.49$$

The ANOVA table

| Source | Sum of Squares | Degrees of freedom | Mean Squares | F ratio |
|---|---|---|---|---|
| Factor A (time) | 51.56 | $(I$-1)=7 | 7.366 | $F_A = 69.49$ |
| Factor B (location) | 0.04 | $(J$-1)=2 | 0.02 | $F_B = 0.19$ |
| Error | 1.49 | $(I$-1)$(J$-1)=14 | 0.106 | |
| Total | 53.09 | $IJ$-1=23 | | |

From the table, we can get following results:

- Since $F_{.05,7,14} = 2.76 < F_A = 69.49$, we reject the null hypothesis and conclude the average amount of the temperature does vary with time.
- Since $F_{.05,2,14} = 3.74 > F_B = 0.19$, we accept the null hypothesis and conclude the average amount of the temperature does not vary with these three cities.